

## **Historic, Archive Document**

Do not assume content reflects current scientific knowledge, policies, or practices.



United States  
Department of  
Agriculture

National  
Agricultural  
Statistics  
Service



Research Division

RD Research Report  
Number RD-98-02

March 1998

# Evaluation of Phase 1 Screening For The 1996 Agricultural Resource Management Study

Mark Apodaca  
Jaki Stanley

103 DEC 15 P-143



**EVALUATION OF PHASE 1 SCREENING FOR THE 1996 AGRICULTURAL RESOURCE MANAGEMENT STUDY.** By Mark Apodaca and Jaki Stanley, Survey Quality Research Section, Survey Research Branch, Research Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Fairfax Virginia 22030. March 1998. Staff Report Number SRB-98-02.

## ABSTRACT

The Agricultural Resource Management Study (ARMS) identifies production practices farm operators use to produce certain commodities, determines the costs to produce those commodities, and describes the economic health of a cross-section of ALL farm and ranch operations. The ARMS design integrated and replaced the Cropping Practices Survey (CPS) and the Farm Costs and Returns Survey (FCRS). The sampling design consists of three phases; a screening phase conducted in early summer, the production practice phase conducted in late summer to early fall and the economic costs and returns phase conducted in late winter and spring.

The objective of the screening phase (P1) is to improve the quality of estimates by increasing the number of usable reports (versus sampling from list frame control data only) without increasing the number of field contacts. The screening phase identifies in-business sample units and operations having targeted commodities. State Statistical Offices during P1 had the option of contacting operations via a telephone interview or using other sources of information to gather commodity specific information about the sampled operation. The presence or absence of a commodity determined whether the operation was eligible to be selected for a Phase 2 (P2) or Phase 3 (P3) interview. This report examines the effectiveness of P1 screening and the efficiency of other screening sources relative to the number of usable reports in P2 and P3.

## KEY WORDS

Agriculture Research Management Study (ARMS); Screening Phase; Screening Sources; Chi-Square; Data Warehouse.

This paper was prepared for limited distribution to the research community outside the U.S. Department of Agriculture.

## ACKNOWLEDGMENTS

The authors would like to thank the ARMS survey team, especially Carl Scott for his technical knowledge and support during our analysis of the ARMS data.

## TABLE OF CONTENTS

ABSTRACT . . . . .	i
SUMMARY . . . . .	iii
INTRODUCTION . . . . .	1
RESULTS AND DISCUSSION . . . . .	2
Screened vs Unscreened Records . . . . .	2
P1 Contacts versus Non-Contacts (Other Screening Sources and Unscreened) . . . . .	3
Efficiency of P1 Questionnaire versus All Other Screening Sources . . . . .	3
Efficiency of Individual Screening Sources . . . . .	5
Efficiency of Screening Sources by ARMS Version . . . . .	6
Response Rates By Respondent Type . . . . .	7
RECOMMENDATIONS . . . . .	9
REFERENCES . . . . .	11
APPENDIX A . . . . .	12
APPENDIX B . . . . .	16

## SUMMARY

The implementation of a screening phase into the Agricultural Resource Management Study (ARMS) was designed to reduce the percentage of out of scope records and increase the number of usable reports while reducing the overall number of contacts made in phase 2 and phase 3 (P2 and P3). A similar sample design used in the 1994 Vegetable Chemical Use Survey yielded a significant reduction in the percentage of records coded out of scope and decreased the cost per usable report (Amrhein 1996). Because of the integration of the Farm Costs and Returns Survey and the Cropping Practices Survey into the ARMS (with a new multi-phase data collection design), no direct cost comparisons can be made to previous years surveys. However, results from the Vegetable Chemical Use Survey suggest that pre-screening operations by telephone rather than selecting a sample directly from the List Frame significantly reduces the number of records coded out of scope, increases the number of usable reports and decreases the number of field contacts made.

The samples for P2 and P3 Version 1 included **screened** and **unscreened** records. Comparisons of the two groups based on the usability of records in P2 and P3 were made. The percentage of out of scope records was significantly higher for the non-screened portions of the P2 and P3 samples.

Data were collected for screening using either a P1 screening questionnaire (designed specifically for the ARMS), which were primarily telephone **contacts**, (with some limited mail and face to face contacts) or data from **other sources**. The majority of the P2 and P3 samples were screened using the P1 screening questionnaire; data from the June Area and List Surveys represented the majority of the other sources used. The ARMS Survey team was concerned that use of sources that were not specifically designed to screen for the ARMS would lead to a higher rate of out of scope records in P2 and P3. Collectively, other sources did have a higher screenout rate. However, there were individual sources that were more efficient than P1 screening contacts. These sources included data from the June Area and Potato Disposition Surveys for P2 screening and data from the March, December, January and June Agricultural Surveys for P3 version 1 screening. Three sources, List Sampling Frame Criteria work, the Acreage and Production Survey and data from the List Frame along with Enumerator Notes were NOT effective, with screenout rates of 20% or greater. These sources are not recommended for use as screening sources in future ARMS surveys.

Analysis of response rates in P2 and P3 were compared for the contact and non-contact portion of each sample. Contacts included records where a P1 contact was made with the operation, and non-contacts included other screening sources and non-screened records. A possible consequence of the addition of a screening phase is that the added burden of the P1 interview may result in a higher percentage of refusals in follow on phases. This was the case for P3 version 1 where the refusal rate was significantly higher for the contact portion of the sample. However the opposite was true in P2; contacts had a significantly lower refusal rate than non-contacts.

The respondent in P1, 2, and 3 was identified as the operator or someone else. P1 screening contacts with respondents other than the operator did not result in an appreciably higher proportion of out of scope records. However, the respondent in the follow on phase is very important. There is a considerably higher proportion of out of scope records in the follow on phases when the follow on respondent is not the Operator. This is true regardless of who the P1 respondent was.

Recommendations include continuation of screening where good control data are not available, using only selected sources of data as alternatives to screening, exploring the use of other resources to screen (such as the forthcoming Data Warehouse) and instructing interviewers to contact operators before coding a record out of scope in Phase 2 or 3.

## INTRODUCTION

The Agricultural Resource Management Study (ARMS) integrated and replaced the Cropping Practices Survey (CPS) and the Farm Costs and Returns Survey (FCRS). The ARMS is being done in cooperation with the Economic Research Service (ERS). The goal of integrating these surveys is to enhance the ERS and NASS estimation programs, as well as the ERS research program in the area of environmental and economic data, and to increase response rates.

The Study measures numerous items of current interest to the agricultural industry, including:

- ▶ The levels of fertilizer and pesticide usage;
- ▶ The adoption of pest management practices;
- ▶ Number of farm operations adopting new technologies;
- ▶ Cost of producing various crop and livestock commodities;
- ▶ Characteristics of high and low cost of production farms;
- ▶ The overall financial situation of farm businesses and households.

The study design is a three stage process. P1 involves screening farm operations to determine if they are still in business and whether they produced commodities of interest in the survey year. After the initial screening phase, all operations that are in business and have the desired commodities are designated as the target population for P2. All operations that are in business and meet minimal sales potential are designated as the target population for P3. All refusals and inaccessibles during P1 are not eligible for selection in P2 and P3. A similar design

was used in the 1994 Vegetable Chemical Use Survey because of a high rate of out of scope records encountered during the 1992 Survey. Comparison between the two years showed that the addition of the screening phase increased the number of usable positive reports in the follow on phase and decreased the cost per usable report (Amrhein 1996). Based on these results a screening phase was added to the ARMS Survey design.

A sequential sampling scheme without replacement is used to select the P2 and P3 samples. This scheme assures that operations will be selected for only one of the eleven versions. During P2 of the study, separate versions of the questionnaire were used to collect crop production practice information from operators growing corn, cotton, potatoes, soybeans, or wheat. Questionnaire versions also collected production practices and cost information for corn, tobacco, and cow/calf operations. P3 of the Study concentrates on overall farm financial and management practices information. All types of farms and ranches are included in the P3 version 1 sample.

State Statistical Offices had the option of contacting operations using the P1 screening questionnaire or other sources of information to gather commodity specific information for selected operations. Other sources of information included data from the List Sampling Frame as well as data from other NASS Surveys. The cost associated with using other sources is minimal. If easily accessible, these data could be gathered to screen the P1 sample. If other sources were not used, information about the operation would usually be obtained through a telephone interview using the P1 screening questionnaire. Additional survey costs are associated with the latter as well as the

burden of an additional contact made with the operation. With response rates steadily declining and increased budget constraints, the use of other sources (when available) may be justified if they are as effective as a P1 contact. In addition, with the ongoing development of the NASS Data Warehouse, more information about a specific operation will be accessible to aid in the screening process.

## RESULTS AND DISCUSSION

### Screened vs Non-screened Records

The final samples for the follow on phases did include records that were not screened. These records included special state funded initiatives and 2500 records added for the Natural Resources Conservation Service. In

addition 14 states were not part of the P1 and P2 survey design. The above records were categorized as non screened records.

The usability of each record in the follow on phases were categorized into three groups: Group 1 contained all completed interviews, Group 2 included out of scope records and Group 3 consisted of refusals and inaccessibles (including 94 records coded with other special response codes). Out of scope records included screenouts (operations that are no longer in business or were misclassified and/or duplicated on the List Frame) and no target commodity records (operations that were still in business but did not have the crop of interest). Table 1 contains the usability status for P2 and P3 Version 1 for the screened and unscreened portion of their sample.

*Table 1. Screened versus Unscreened Records by Usability of P2 and P3 Records.*

n (Row %)	Screened in P1 (P1 contacts or other sources)	Response Summary			Total
		Completed Interviews	Out of Scope	Refusal/ Innac	
P2 (v2-v11)	Yes	7,722 (70.49)	630 (5.75)	2,603 (23.76)	10,955 (72.65)
	No	1,910 (46.30)	1,214 (29.43)	1,001 (24.27)	4,125 (27.35)
	Total	9,632 (63.87)	1,844 (12.23)	3,604 (23.90)	15,080
P3 (v1)	Yes	6,855 (53.52)	591 (4.61)	5,363 (41.87)	12,809 (91.43)
	No	659 (54.87)	101 (8.41)	441 (36.72)	1,201 (8.57)
	Total	7,514 (53.63)	692 (4.94)	4,238 (41.43)	14,010

P2:  $\chi^2$  (2,n=15,080)=1649, p < .0001

P3:  $\chi^2$  (2,n=14,010)= 48, p < .0001

A total of 15,080 records were analyzed across all eleven versions of P2. Overall, 72.7% of the sample selected for P2 were screened in P1. In P3, 91.4% of the sample was screened. Chi-Square tests were performed to test whether the usability of a record in P2 and P3 was independent of whether it was screened or unscreened. Both tests were highly significant for the two follow on phases. The percentage of out of scope records was significantly higher for the non-screened portion of the P2 sample at 29.4% versus 5.75%. Excluding the 2500 non-screened NRCS samples (which will not be included in planned future ARMS) the percentage of out of scope records drops to 12.74% which is still significantly higher than screened records [ $\chi^2$  (2,N=12,580) = 240, p < .0001]. In P3 the percentage of out of scope records was 4.61% and 8.41% for the screened and unscreened portion of the sample.

(The percentage of records classified as out of scope is much higher in P2 because a record had to both be in business and have the commodity of interest. P3 Version 1 collected general farm costs and returns data and did not target specific commodities.)

The results above suggest that a screening phase in the ARMS survey design reduced the number of records coded out of scope in the follow-on phases. In P2 the percentage of usable records was significantly higher for the screened portion of the sample than the unscreened portion. In addition, a comparison between the P1 screened records and their list frame control data also shows the effectiveness of screening (Scott, 1997). Records with zero control data that reported the commodity of interest during screening ranged from 2% (for flue cured tobacco) to a high of 28% (for Beef Cows (NAHMS)). These records would not have been eligible

for sample selection if list frame control data was the sole criteria used. Also, the percent of records that reported zero commodity data during screening but had positive control data ranged from 11.5% (Corn PPR) to 36% (Tobacco PPCR). If sampling directly from the list frame based on control data these operations would have been contacted with costly personal interviews only to be classified as out of scope and non-usable.

#### P1 Contacts Versus Non-Contacts (Other Screening Sources and Unscreened)

Refusal and completion rates were compared between P1 screening questionnaires (contacts) and non-contacts. Non-contacts included records screened by other sources and non-screened records. A possible consequence of the addition of a screening phase is that the added burden of the screening interview may result in a higher percentage of refusals in follow on phases. Although the average interview time in P1 was only 5 minutes, an additional contact was made with the operation thus potentially increasing the probability that the operation may refuse when contacted again in the follow on phase. However, the conclusions were different for the two follow on phases. In P2, the refusal rate for P1 contacts was significantly less at 18.5% versus 25.6% for non-contacts [ $\chi^2$  (1,N=15,080) = 90, p < .0001]. The opposite was true in P3, contacts had a higher refusal rate than non-contacts at 38.2% compared to 34.2% for non-contacts. [ $\chi^2$  (1,N=14,010) = 18, p < .0001]

#### Efficiency of P1 Questionnaire vs All Other Screening Sources

Data collection for the P1 screening phase included mail, telephone and field interviews. In addition, State Statistical

Offices were allowed to use other sources of information to gather commodity specific information as an alternative to a P1 contact. The total NASDA cost associated with P1 screening of 49,000 records was approximately 172,000 dollars or about \$3.48 per screened record (Rutz, 1997). Information from other sources was gathered by state office personnel and no additional NASDA costs are associated with this screening method. If other sources are as efficient in screening as a contact, the use of other sources as a screening technique would be a cost effective alternative.

Screened records in P2 and P3 were categorized into two groups. Group 1 consisted of records screened using a P1 screening questionnaire and group 2 contained records screened using information from other sources. A total of 10,955 records or 72.6% of the P2 sample were screened. Of the total records screened for

P2, 88.4% were screened using a P1 questionnaire. For P3 Version I, 91.4% of the records were screened. The percentage of records screened using other sources was somewhat higher in P3 at 27.2% (only 11.6% used other sources in P2). See Table 2.

Chi-Square tests indicated that the usability of records in the follow on phases are not independent of the method of screening used. Other sources resulted in a higher percentage of records found out of scope in the follow on phases. This did not appear to be related to the age of the control data. (The average age of control data for completed interviews was only 6 months less than control data for screenouts and operations with no target commodity.) Although collectively other sources had a higher percentage of out of scope records, there were individual sources which had lower percentages of records coded out of scope.

*Table 2. Screening Methods versus Usability of P2 and P3 records.*

n (Row %)	P1 Screening Method	Response Summary			Total
		Completed Interviews	Out of Scope	Refusal/ Innac	
P2	P1 Questionnaire	6926 (71.48)	488 (5.04)	2275 (23.48)	9,689 (88.44)
	Other Sources	796 (62.88)	142 (11.22)	328 (25.91)	1,266 (11.56)
	Total	7722 (70.49)	630 (5.75)	2603 (23.76)	10955
P3 (v1)	P1 Questionnaire	5003 (53.63)	365 (3.91)	3961 (42.46)	9329 (72.83)
	Other Sources	1852 (53.22)	226 (6.49)	1402 (40.29)	3480 (27.17)
	Total	6855 (53.52)	591 (4.61)	5363 (41.87)	12809

P2:  $\chi^2$  (2, n=10955) = 89, p < .0001

P3:  $\chi^2$  (2, n=12809) = 40, p < .0001

### Efficiency of Individual Screening Sources

The number of records screened, percent usable and percent out of scope for each source are shown in Table 3 and Table 4 (for sources used to screen at least 10 records). P1 contacts accounted for 88.4% of all records screened in P2. Information from the June List and Area Surveys, at 3.1% were the most commonly used other sources. June Area was most efficient among the three sources with only 3.8% of the records coded out of scope, P1 contacts and data from the June List Survey had screenout rates of 5.0% and 6.1%. The most efficient screening source, data from the Potato Disposition Survey had a 0.0% screenout rate (57 screened).

Three sources were not efficient methods and had out of scope rates of 20% or greater. Information from LSF Criteria work,

Acreage and Production Surveys, and List Sampling Frame along with Enumerator Notes had screenout rates of 33.3% (18 Screened), 23.2% (293 Screened) and 23.1% (121 screened) respectively. These sources are NOT recommended for use as a screening source in future surveys.

In P3, 91.4% of the Version 1 sample was screened. (Version 1 of P3 is not commodity specific, so operations are only screened to determine whether or not they were in business.) P1 contacts accounted for 72.8% of the screening and had an out of scope rate of 3.9%. As in P2, June Area and June List information were the second and third most commonly used screening sources. June Area and List data were good sources of screening information with out of scope rates of 7.0% for June Area and 2.1% for June List.

*Table 3. P2 Inefficiency rating for Individual Screening Sources.  
(Minimum of 10 screened).*

Screening Sources	Total	Completed Interviews	% usable	Out of Scope	% Out of Scope
Potato Disposition	57	36	63.2	0	0.0
June Area	341	248	72.8	13	3.8
<b>P1 Contact</b>	<b>9,689</b>	<b>6,926</b>	<b>71.5</b>	<b>488</b>	<b>5.0</b>
June List	344	229	66.6	21	6.1
Potato Stocks	15	7	46.7	1	6.7
December Ag	29	13	44.8	2	6.9
March Ag	23	15	65.2	2	8.7
LSF, Enum Notes	121	58	48.0	28	23.1
A & P	293	163	55.6	68	23.2
Not Screened	4125	1910	46.3	1214	29.4
LSF Criteria	18	10	55.5	6	33.3

*Table 4. P3 Version I, Inefficiency rating for Screening Sources.  
( Minimum 10 screened).*

Screening Source	Total Screened	Completed Interviews	% Usable	Out of Scope	% Out of Scope
March Ag	13	9	69.2	0	0.0
December Ag	15	7	46.7	0	0.0
January Cattle	28	20	71.4	0	0.0
July Cattle	37	15	40.5	6	1.8
June List	291	136	46.7	6	2.1
<b>P1 Contact</b>	<b>9329</b>	<b>5003</b>	<b>53.6</b>	<b>365</b>	<b>3.9</b>
A & P	53	37	69.8	3	5.7
June Area	2947	1570	53.2	205	7.0
Not Screened	1201	659	54.9	101	8.4
LSF Criteria	30	18	60.0	3	10.0
LSF, Enum Notes	48	26	54.2	8	16.7

There were several other good sources (at least 10 screened) that had screenout rates of 0.0%, these included data from the March, December, January and July Agricultural Surveys. As in P2, LSF Criteria and LSF information along with enumerator notes once again rated poorly with screenout rates of 16.7% and 10.0%.

#### Efficiency of Screening Sources by ARMS Version

The percentage of records coded out of scope by ARMS version varied considerably by screening method (not screened, screened in Phase 1 with a questionnaire or screened using some other source) (Table 5). Appendix A contains the screening sources by version for P2 and P3. The Wheat Production Practices Report (P2, Version 7) had an out of scope rate of 54.8% for non-

screened records. Screened records did not fare much better with 24.3% (other sources) and 10.0% (P1 contacts) of their records coded out of scope. The high percentage of records coded out of scope for the screened wheat sample was a result of P1 records being screened on planted acres rather than on harvested acres. A large amount of wheat acreage was lost in 1996 leading to high out of scope rates.

The use of other sources to screen for the Flue Cured Tobacco Production Practices and Cost Report were also not very efficient. LSF information had an out of scope rate of 46.6% with information from the June Ag List following closely behind at 37.5%. LSF information rated poorly because data maintained on the List Frame is an aggregate value not only for flue-cured, but for all varieties of tobacco.

Table 5. Percentage of Out of Scope by Screening Method and ARMS Version.

Arms Version	Screening Methods					
	# of Records			% Out of Scope		
	Not Screened	P1 Q's	Other Source	Not Screened	P1 Q's	Other Source
<b>P2</b>						
Corn	0	1850	239	n/a	4.6	12.5
Tobacco	0	368	112	n/a	7.6	24.1
Cow/Calf	0	1586	174	n/a	7.7	9.2
Corn	84	582	33	53.6	2.9	0.0
Soybeans	0	1251	80	n/a	4.0	1.3
Wheat	504	908	210	54.8	10.0	24.3
Cotton	158	939	94	8.9	3.7	7.4
Potato	8	440	203	37.5	4.8	2.5
Multi-Crop	3371	1184	97	26.0	2.4	3.1
State Projects	0	581	24	n/a	1.4	8.3
<b>P3</b>						
General	1201	9329	3480	8.4	4.0	6.5
Corn	0	1300	159	n/a	0.8	2.5
Tobacco	0	274	77	n/a	3.6	2.6
Cow/Calf	0	1320	55	n/a	0.7	0.0

Other sources used for the Corn (version 5), Soybeans (version 6) and the Potato Production Practices (version 8) were more efficient than P1 contacts. June Agricultural Survey data were the most commonly used screening source for the Corn and Soybean P2 Versions. Information from the Potato Disposition Survey had zero records coded out of scope in Version 8. Information from NASS Surveys like these which collect commodity specific information such as the Potato Disposition Survey as well as data from the June List and Area Surveys are

screening sources that should be considered as substitutes for a P1 contact.

#### Response Rates by Respondent Type

The respondent in each P1, P2, and P3 was identified as the operator or someone else. The following four groups were created (P1 Respondent-Follow on Phase Respondent): Operator-Operator, Other-Operator, Operator-Other and Other-Other. Usability of P2 and P3 records were not independent across the four respondent groups [P2:  $\chi^2$  (6, n=8608) = 426,  $p < .0001$ ; P3:  $\chi^2$  (6,

$n=7561) = 144$ ,  $p < .0001$ ]. The percentage of out of scope records is significantly higher when someone other than the operator is contacted in P2 or P3. This is only true for the follow on phase respondent, P1 respondents other than the operator did not result in appreciably higher

percentage of out of scope records in the follow on phases. It should be stressed to enumerators that an operation should not be coded out of scope unless they have made every attempt to speak to the operator in the P2 and P3.

*Table 6. Response Summary by P1 and P2 Respondents.*

n Expected (Row %)	Respondents (P 1- P2)	Response Summary			Total
		Completed Interviews	Out of Scope	Refusal/ Innac	
P2	Operator- Operator	5587 5393 (83.25)	233 284 (3.47)	891 1034 (13.27)	6711 (77.96)
	Other- Operator	1066 1085 (78.96)	51 57 (3.78)	233 208 (17.26)	1350 (15.68)
	Operator- Other	177 310 (45.97)	62 16 (16.10)	146 59 (37.92)	385 (4.47)
	Other- Other	87 130 (53.70)	19 7 (11.73)	56 25 (34.57)	162 (1.88)
	<b>Total</b>	<b>6917 (80.36)</b>	<b>365 (4.24)</b>	<b>1326 (15.40)</b>	<b>8608</b>

Table 7. Response Summary by P1 and P3 Respondents

n Expected (Row %)	Respondents (P1-P3)	Response Summary			Total
		Completed Interviews	Out of Scope	Refusal/ Inacc	
P3 (v1)	Operator- Operator	3921 3805 (69.60)	93 124 (1.65)	1620 1704 (28.75)	5634 (74.51)
	Other- Operator	782 865 (61.05)	22 28 (1.72)	477 388 (37.24)	1281 (16.94)
	Operator- Other	254 272 (63.18)	33 9 (8.21)	115 122 (28.61)	402 (5.32)
	Other- Other	150 165 (61.48)	18 5 (7.38)	76 74 (31.15)	244 (3.23)
	Total	5107 (67.54)	166 (2.20)	2288 (30.26)	7561

## DISCUSSION AND RECOMMENDATIONS

The introduction of the screening phase into the ARMS was designed to increase the number of usable reports without increasing the number of field contacts. The loss of sampling units, due to operations refusing or being coded out of scope in costly follow on personal contacts, impact the survey indications and survey costs. Through screening, prior information about the operation can be obtained in less expensive telephone or mail contacts or from other sources, thus increasing the chances that the operation will be in scope when contacted in person in follow on phases. (However, the exclusion of Phase 1 refusals from follow on phases has the potential to seriously bias follow on phase estimates. Currently, there is no way to know whether or how Phase 1 non-respondents differ from respondents.) Results from the 1996 ARMS suggest that the screening phase of the sample design did decrease the number of records coded out of

scope in the follow on phases. The percentage of records coded out of scope was considerably higher for non-screened records.

The effectiveness of screening depends on the accuracy of the data gathered about the sampled operation. There were some sources such as June Area and commodity specific surveys (Potato disposition) that were as efficient or better than a P1 contact and less costly.

Based on the analysis of the 1996 ARMS data, the continued use of a screening phase is recommended. In order to maximize the full extent of the screening phase, increasing the number of usable records, cutting total survey costs and minimizing the number of personal contacts made with non-useable sampling units, the following recommendations are offered:

## Recommendations

1. Continue to use screening for ARMS or other surveys involving personal interviews where good control data are not available. For example, control data would not be considered good in situations where crops are likely to change from year to year (as with vegetables), or when the source of data is unreliable (as was the case with NRCS samples) or does not match the survey concepts closely (i.e. control data is for ALL tobacco, survey targets flue cured tobacco only). The number of usable reports in follow on phases are increased compared to a selecting a sample directly from the List Frame. Information obtained in P1 is more accurate than List Frame data thus leading to fewer records coded out of scope in the follow on phases. However, a careful evaluation of the relative costs of adding a screening phase versus increasing the Phase 2/3 samples sizes should be conducted. If significant portions of the P1 screening phase sample do not screen out, adding a screening phase is probably NOT cost effective. See Appendix B for a discussion about how cost benefits of screening might be calculated.
2. Information from the June Survey as well as data from other NASS commodity specific surveys such as the Potato Disposition Survey and the July Cattle Surveys are an efficient alternative to a P1 contact. These surveys are conducted relatively close to the time P1

contacts are made and collect specific data similar to the P1 screening data. Because an attempt is made to limit the number of operations that are selected for participation in both ARMS and other surveys the number of sample units with other sources of screening information will likely be fairly small. However, in those cases where these data are available the efficient alternative sources should be used to screen operations instead of a P1 contact.

For the 1996 ARMS, the List Sampling Frame, List Frame Criteria work, and the Acreage and Production Surveys were not effective screening sources. Therefore, these sources are NOT recommended as screening sources in future ARMS screening. Weaknesses of these sources include the lack of a current year contact and non-probability data collection procedures, and collection of data that does not closely match the P1 screening data.

3. Develop a model that uses the NASS Data Warehouse as a screening source for the P1 sample. A Data Warehouse will contain a wealth of historical data for use to screen an operation. Once the P1 sample is selected a match with historical data can be made and a profile of the operation can be generated. Information from this profile may be used to determine if a contact is necessary to obtain more information in P1. As was shown in this analysis, some

sources and types of information are better for screening purposes than others. Which data contained in a Data Warehouse will be effective for screening purposes should be further explored.

4. Continue stressing to enumerators that every effort should be made to

talk to the primary operator prior to coding a record out of scope in P2 or 3. The proportion of out of scope records was significantly higher when the follow on phase respondent was not the operator.

## REFERENCES

Amrhein, J (1996), Evaluation of the 1994 Vegetable Chemical Use Survey Screening Phase", NASS Staff Report Number SSB-96-01, United States Department of Agriculture, National Agricultural Statistics Service.

Scott, C. (1997), "Highlights of P1 Screening", Inter-office Memo To Agricultural Resource Management Team, United States Department of Agriculture, National Agricultural Statistics Service.

Rutz, J. (1997), "P1 and 2 NASDA Cost Analysis-Final", Inter Office Memo To Agricultural Resource Management Team, United States Department of Agriculture, National Agricultural Statistics Service.

"Agriculture Resources Management Study Administration Manual" (1997), United States Department of Agriculture, National Agricultural Statistics Service.



## APPENDIX A:

### Usability of P2 Records By Screening Source in P1

#### Sorted By Most Reliable Screening Source

----- ARMS SURVEY VERSION #=Corn PPCR (2) -----

Screening Source	Complete	Refusals/Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag List	19	14	1	34	55.8824	2.9412
Contacted	1246	518	86	1850	67.3514	4.6486
Jun Ag Area	85	38	6	129	65.8915	4.6512
March ag	4	3	1	8	50.0000	12.5000
Dec Ag	3	2	1	6	50.0000	16.6667
A & P	14	1	4	19	73.6842	21.0526
LSF, Enum, Etc	18	3	12	33	54.5455	36.3636
LSF Criteria	4	1	5	10	40.0000	50.0000
<hr/>						
<b>TOTAL</b>	<b>1393</b>		<b>116</b>	<b>2089</b>		

----- ARMS SURVEY VERSION #=Tob PPCR (3) -----

Screening Source	Complete	Refusals/Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jan cattle	1	2	0	3	33.333	0.0000
Sept Ag	2	0	0	2	100.000	0.0000
Contacted	252	88	28	368	68.478	7.6087
Dec Ag	6	3	1	10	60.000	10.0000
March ag	6	2	1	9	66.667	11.1111
Jun Ag Area	6	1	1	8	75.000	12.5000
LSF Criteria	6	1	1	8	75.000	12.5000
A & P	21	7	6	34	61.765	17.6471
Jun Ag List	2	3	3	8	25.000	37.5000
LSF, Enum, Etc	14	2	14	30	46.667	46.6667
<hr/>						
<b>TOTAL</b>	<b>316</b>		<b>55</b>	<b>480</b>		

----- ARMS SURVEY VERSION #=Cow/Calf PPCR (4) -----

Screening Source	Complete	Refusals/Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag Area	35	5	0	40	87.500	0.0000
July Cattle	4	2	0	6	66.667	0.0000
Jan sheep	1	0	0	1	100.000	0.0000
Contacted	1158	305	123	1586	73.014	7.7554
Jun Ag List	13	2	2	17	76.471	11.7647
A & P	68	28	14	110	61.818	12.7273
<hr/>						
<b>TOTAL</b>	<b>1279</b>		<b>139</b>	<b>1760</b>		

Usability of P2 Records  
By Screening Source in P1

Sorted By Most Reliable Screening Source

----- ARMS SURVEY VERSION #=Corn PPR (5) -----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag List	8	7	0	15	53.333	0.0000
Jun Ag Area	11	4	0	15	73.333	0.0000
Jan cattle	0	1	0	1	0.000	0.0000
LSF, Enum, Etc	2	0	0	2	100.000	0.0000
Contacted	414	151	17	582	71.134	2.9210
Not Screened	26	13	45	34	30.952	53.5714
<b>TOTAL</b>	<b>461</b>		<b>62</b>	<b>529</b>		

----- ARMS SURVEY VERSION #=Soy R (6) -----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag Area	24	6	0	30	80.0000	0.00000
Jun Ag List	43	6	1	50	86.0000	2.00000
Contacted	934	267	50	1251	74.6603	3.99680
<b>TOTAL</b>	<b>1001</b>		<b>51</b>	<b>1331</b>		

----- ARMS SURVEY VERSION #=Wheat PPR (7)-----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
March ag	1	0	0	1	100.000	0.0000
LSF, Enum, Etc	3	0	0	3	100.000	0.0000
Jun Ag List	36	10	3	49	73.469	6.1224
Contacted	640	177	91	908	70.485	10.0220
July Ag Yld	6	1	1	8	75.000	12.5000
Jun Ag Area	11	5	3	19	57.895	15.7895
A & P	60	26	44	130	46.154	33.8462
Not Screened	166	62	276	504	32.937	54.7619
<b>TOTAL</b>	<b>923</b>		<b>418</b>	<b>1622</b>		

Usability of P2 Records  
By Screening Source in P1

Sorted By Most Reliable Screening Source

----- ARMS SURVEY VERSION #="Cotton PPR (8)-----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
March ag	1	0	0	1	100.000	0.0000
LSF, Enum, Etc	1	0	0	1	100.000	0.0000
Jun Ag Area	29	6	1	36	80.556	2.7778
Contacted	659	245	35	939	70.181	3.7274
Not Screened	92	52	14	158	58.228	8.8608
Jun Ag List	33	17	6	56	58.929	10.7143
<b>TOTAL</b>	<b>815</b>		<b>56</b>	<b>1191</b>		

----- ARMS SURVEY VERSION #="Potato PPR (9)-----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag Area	9	2	0	11	81.818	0.0000
March ag	3	1	0	4	75.000	0.0000
Dec Ag	4	9	0	13	30.769	0.0000
Potato Disp	36	21	0	57	63.158	0.0000
May Ag Yld	1	0	0	1	100.000	0.0000
Jun Ag List	33	20	2	55	60.000	3.6364
LSF, Enum, Etc	17	28	2	47	36.170	4.2553
Contacted	290	129	21	440	65.909	4.7727
Potato Stks	7	7	1	15	46.667	6.6667
Not Screened	1	4	3	8	12.500	37.5000
<b>TOTAL</b>	<b>401</b>		<b>29</b>	<b>651</b>		

----- ARMS SURVEY VERSION #="Crop PPR (10)-----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag Area	26	11	0	37	70.270	0.0000
July Cattle	1	0	0	1	100.000	0.0000
Wheat OBY	1	0	0	1	100.000	0.0000
LSF, Enum, Etc	1	1	0	2	50.000	0.0000
Contacted	827	328	29	1184	69.848	2.4493
Jun Ag List	38	15	3	56	67.857	5.3571
Not Screened	1625	870	876	3371	48.205	25.9864
<b>TOTAL</b>	<b>2519</b>		<b>908</b>	<b>4652</b>		

**Usability of P2 Records  
By Screening Source in P1**

**Sorted By Most Reliable Screening Source**

----- ARMS SURVEY VERSION #=State Projects-----

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
Jun Ag List	4	0	0	4	100.000	0.0000
July Cattle	0	1	0	1	0.000	0.0000
LSF, Enum, Etc	2	1	0	3	66.667	0.0000
Contacted	506	67	8	581	87.091	1.3769
Jun Ag Area	12	2	2	16	75.000	12.5000
-----						
TOTAL	524		10	605		
=====						
	9632		1844	15080		

**Usability of P3 Version 1 Records  
By Screening Source in P1**

**Sorted By Most Reliable Screening Source**

Screening Source	Complete	Refusals/ Innac	Out Of Scope	Total	% Usable	% Out of Scope
July Cattle	15	22	0	37	40.541	0.0000
Jan cattle	20	8	0	28	71.429	0.0000
Dec Ag	7	8	0	15	46.667	0.0000
March ag	9	4	0	13	69.231	0.0000
Potato Disp	6	1	0	7	85.714	0.0000
Wheat OBY	2	0	0	2	100.000	0.0000
July Ag Yld	2	0	0	2	100.000	0.0000
Potato Stks	1	0	0	1	100.000	0.0000
FCRS	0	1	0	1	0.000	0.0000
State Funded	1	0	0	1	100.000	0.0000
Jan sheep	1	0	0	1	100.000	0.0000
Jun Ag List	136	149	6	291	46.735	2.0619
Contacted	5003	3961	365	9329	53.628	3.9125
A & P	37	13	3	53	69.811	5.6604
Jun Ag Area	1570	1172	205	2947	53.275	6.9562
Not Screened	659	441	101	1201	54.871	8.4097
LSF Criteria	18	9	3	30	60.000	10.0000
LSF, Enum, Etc	26	14	8	48	54.167	16.6667
Apple disp	1	1	1	3	33.333	33.3333
=====						
	7514		692	14010		

## APPENDIX B. Cost Evaluation of Screening for ARMS

For the 1996 ARMS survey NASDA costs were:

P1: \$172,004; \$3.48 per sample

P2: \$1,337,881; \$76 per sample

The idea behind the screening phase was to increase the number of usable completed reports in the follow on phase (Phase 2 and 3 for commodity specific samples and Phase 3 for the non-commodity specific version). By eliminating screenouts and some other non-response in telephone calls (P1), usable reports (as a percentage of total contacts) were increased in follow on phases in 1996. However, the total number of usable reports could also be increased simply by increasing the follow-on phase initial sample size. Is screening cheaper than increasing the follow on sample size?

Some of the sample units in P2 were not screened during P1 (these included samples added for state initiatives, and states that did not conduct screening). The P2 screenout rate was ~13% for these samples compared to ~5% for samples that had been screened. Therefore we would expect an additional 8% of the sample to screenout if P1 were eliminated.

In P1 there was also a refusal rate of ~12%. These sample units were not eligible for P2 screening. If we assume that these people are hard core refusals and all would have refused in P2, we would expect this additional 12% refusals in P2 if P1 screening were eliminated. (This may be a generous assumption, since response rates are typically much higher for personal interviews than for telephone interviews.)

In order to arrive at the same number of usable reports as obtained in 1996 P2, these additional 8% and 12% of the sample would have to be added to P2. If we assume each non-usable personal contact costs somewhat less than the overall costs per sample (no actual interview will be conducted) we can multiply the number of sample units in this extra 20% by \$40 (estimated cost for personal screenout, or non-response) to arrive at additional costs. If this cost is less than the cost of the P1 screening it is cheaper to add additional samples in the follow-on phase and eliminate the P1 screening entirely.

If control data is much worse than information obtained during screening, screening becomes much more cost effective. This would most likely be the case for commodities that may vary from year to year (for example specific vegetables) or for commodities where the control data does not exactly match the screening criteria.

For example, tobacco control data is for ALL tobacco, the 1996 ARMS was for flue cured tobacco only. For the 1996 ARMS P2 Tobacco, there were 316 usable reports, from a sample size of 480. Of the records screened in P1, 36% reported 0 acres but had positive control data. If P1 was eliminated, we would expect these 36% to screenout in P2. We would also still expect to get the same 12% refusal increase from P1. Therefore, to obtain the same number of usable reports, we would have to increase the sample size by 36% + 12% or 443 records. Multiplied by the cost to contact these operations in person (\$40) this would cost \$17720 beyond the current

P2 NASDA costs. This would be contrasted with screening 1440 (approximately 3 times the sample) at \$3.48/sample or \$5011. In this case the cost of screening would clearly be cheaper than contacting these people in person.

The contrasting case to this is screening for an item for which we have considerably better control data, or screening for much more general information like business status. When control data is good, it can effectively replace the screening phase and has no associated NASDA costs. For example, for the Cotton PPR there were 815 usable reports from an initial P2 sample of 1191. Only 4.3% of the P1 records who had positive control data reported that they did not have cotton. These people would have screened out in P2 if P1 were eliminated. Again, we would still expect an increase of 12% in refusals if P1 were eliminated. Therefore, in this case we would have to add 16% more records to this sample. The cost of these additional records (at \$40 per record) would probably be cheaper than the cost of telephoning the large Phase 1 sample (at \$3.48 per record).

In 1996, screening was also conducted for sample units included only in P3 which were not targeted at a specific commodity. For this ARMS subsample, the screening was only used to determine whether the operation was in or out of business. We would expect that our control data would be much better for determining business status than the presence of any particular crop during the previous year. Therefore, screening this entire subsample is probably less cost effective than simply adding a small number to the follow-on phase sample to compensate for out of business operations.

Overall, screening is effective only if the number screened considerably reduces the non-usable contacts and the number of operations screened is not excessively high. There are several ways in which screening can be made more cost effective. The simplest way would simply be to screen fewer records. A sample of 15K could easily be selected from a screened sample of only 30K not 49K. This would result in a savings of over \$70K in screening costs. (In addition, this screening need not be done by office telephone enumerators. Field interviewers could be instructed to call their sampled respondents to verify that they have the commodity of interest before they contact them in person.)

Another way to reduce costs is to target screening only to commodities we expect to have poor or somewhat mismatched control data. For those commodities that we expect to have good control data, sample sizes can be increased to account for the minimal expected increases in screenouts and refusals and the screening phase eliminated. For other commodities with poor expected control data or where control data does not match screening specifications, a screening phase could be implemented only for these records.



